

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

THE NEW YORK TIMES COMPANY,

Plaintiff,

v.

MICROSOFT CORPORATION, OPENAI, INC.,
OPENAI LP, OPENAI GP, LLC, OPENAI, LLC,
OPENAI OPCO LLC, OPENAI GLOBAL LLC,
OAI CORPORATION, LLC, OPENAI
HOLDINGS, LLC,

Defendants.

Case No. 1:23-cv-11195 (SHS) (OTW)

**MEMORANDUM OF LAW
IN SUPPORT OF DEFENDANT
OPENAI OPCO, LLC'S
OBJECTION TO
NONDISPOSITIVE ORDER
PURSUANT TO FEDERAL RULE
OF CIVIL PROCEDURE 72(a)**

TABLE OF CONTENTS

	Page
I. INTRODUCTION	1
II. BACKGROUND	2
A. The Times’s Contradictory Allegations and Actions	2
B. OpenAI Hewed Closely to Precedent When Seeking Evidence	5
C. The Order Denied OpenAI’s Request by Ignoring Governing Law	7
III. STANDARD OF REVIEW	9
IV. ARGUMENT	9
A. The Order is Contrary to Binding Fair Use Precedent	10
1. The Order Erroneously Concludes that Plaintiff’s Conduct is Almost Entirely Irrelevant to Fair Use	11
2. The Order Misconstrues the Market Harm Inquiry	13
3. The Order Overlooks the Relevance to Transformativeness and Public Benefit	16
B. The Order Shields Relevant Evidence from Discovery	18
1. The Documents OpenAI Seeks are Directly Relevant	18
2. The Court Relied on an Unwarranted Factual Assumption about the Times’s Own Conduct	19
C. The Order Ignores the Relevance to Damages and Other Remedies	20
V. CONCLUSION	21

TABLE OF AUTHORITIES**Page(s)****CASES**

<i>Am. Geophysical Union v. Texaco Inc.</i> , 60 F.3d 913 (2d Cir. 1994).....	<i>passim</i>
<i>Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith</i> , 598 U.S. 508 (2023).....	8
<i>Arista Records LLC, et al. v. Lime Group LLC</i> , No. 06-cv-5936, Dkt. 363 (S.D.N.Y. Nov. 19, 2010).....	18
<i>Authors Guild v. Google, Inc. (Google Books)</i> , 804 F.3d 202 (2d Cir. 2015).....	5, 11, 16
<i>Bamonte v. Charatan</i> , No. 22-cv-0795, 2023 WL 4201416 (S.D.N.Y. June 27, 2023)	10
<i>Bill Graham Archives v. Dorling Kindersley Ltd.</i> , 448 F.3d 605 (2d Cir. 2006).....	13
<i>Campbell v. Acuff-Rose Music, Inc.</i> , 510 U.S. 569 (1994).....	8, 10, 11, 16
<i>Chen-Oster v. Goldman, Sachs & Co.</i> , 293 F.R.D. 557 (S.D.N.Y. 2013)	9, 17
<i>Compudyne Corp. v. Shane</i> , 244 F.R.D. 282 (S.D.N.Y. 2007)	16
<i>Concord Music Group, Inc. v. Anthropic PBC</i> , No. 24-cv-03811 (N.D. Cal.)	4
<i>Condit v. Dunne</i> , 225 F.R.D. 100 (S.D.N.Y. 2004)	9, 17
<i>Coventry Capital U.S. LLC v. EEA Life Settlements Inc.</i> , 333 F.R.D. 60 (S.D.N.Y. 2019)	9
<i>Daval Steel Prod., a Div. of Francosteel Corp. v. M/V Fakredine</i> , 951 F.2d 1357 (2d Cir. 1991).....	9
<i>Dawe v. Capital One Bank</i> , No. 04-cv-40192, 2005 WL 8176463 (D. Mass. Sept. 27, 2005).....	18, 19

Page(s)

<i>Dubus v. Nvidia Corp.</i> , No. 24-cv-02655 (N.D. Cal.)	4
<i>Forte v. Liquidnet Holdings, Inc.</i> , No. 14-cv-2185, 2014 WL 12775220 (S.D.N.Y. Dec. 2, 2014)	9
<i>Francis v. A & E Stores, Inc.</i> , No. 06-cv-1638, 2008 WL 2588851 (S.D.N.Y. June 26, 2008)	20
<i>Freedom Card, Inc. v. JPMorgan Chase & Co.</i> , 432 F.3d 463 (3d Cir. 2005)	12
<i>Google LLC v. Oracle Am. Inc.</i> , 593 U.S. 1 (2021)	passim
<i>Int’l Bus. Machs. Corp. v. Naganayagam</i> , No. 15-cv-7991, 2016 WL 11483935 (S.D.N.Y. Dec. 9, 2016)	12
<i>J.L. v. Alphabet Inc.</i> , No. 23-cv-03440 (N.D. Cal.)	4
<i>Kadrey v. Meta Platforms</i> , No. 23-cv-03417 (N.D. Cal.)	4
<i>Kirtsaeng v. John Wiley & Sons, Inc.</i> , 579 U.S. 197 (2016)	20
<i>Makkai v. Databricks, Inc.</i> , No. 3:24-cv-02653 (N.D. Cal.)	4
<i>Millette v. Google LLC</i> , No. 24-cv-04708 (N.D. Cal.)	4
<i>Millette v. Nvidia Corp.</i> , No. 24-cv-05157 (N.D. Cal.)	4
<i>Nazemian v. Nvidia Corp.</i> , No. 23-cv-01454 (N.D. Cal.)	4
<i>O’Nan v. Databricks, Inc.</i> , No. 24-cv-01451 (N.D. Cal.)	4
<i>On Davis v. The Gap, Inc.</i> , 246 F.3d 152 (2d Cir. 2001)	16

Page(s)

<i>Pearson Educ., Inc. v. Chegg, Inc.</i> , No. 21-cv-16866, 2023 WL 3775317 (D.N.J. June 2, 2023).....	12, 21
<i>R.F.M.A.S., Inc. v. So</i> , 748 F. Supp. 2d 244 (S.D.N.Y. 2010).....	9
<i>Sega Enterprises Ltd. v. Accolade, Inc.</i> , 977 F.2d 1510 (9th Cir. 1992)	16
<i>Stewart v. Abend</i> , 495 U.S. 207 (1990).....	10
<i>TD Bank N.A. v. Hill</i> , 928 F.3d 259 (3d Cir. 2019).....	20
<i>Zhang v. Google LLC</i> , No. 24-cv-02531 (N.D. Cal.)	4

STATUTES

17 U.S.C.	
§ 107.....	10, 11, 13
§ 504(c)(2)	20
§ 505.....	20

RULES

Fed. R. Civ. Proc.	
26.....	9
72(a)	1, 9, 15

TREATISE

Wright & Miller, 8 FED. PRAC. & PROC. CIV. § 2015	18
---	----

CONSTITUTIONAL PROVISION

U.S. CONST. Article 1, § 8, cl. 8	16
---	----

OTHER AUTHORITIES

Alexandra Berzon, et al., <i>Inside the Movement Behind Trump’s Election Lies</i> , N.Y. Times (Oct. 28, 2024), https://www.nytimes.com/interactive/2024/10/28/us/politics/inside-the-movement-behind-trumps-election-lies.html	2
---	---

Page(s)

Carl Franzen, <i>Marketing Software Anyword Adds Integrations to ChatGPT, Notion AI, and Canva AI</i> , Venture Beat (July 18, 2023), https://venturebeat.com/ai/marketing-software-anyword-adds-integrations-to-chatgpt-notion-ai-and-canva-ai/	2
Jon Gertner, <i>Wikipedia's Moment of Truth</i> , The New York Times Magazine (July 18, 2023), available at https://www.nytimes.com/2023/07/18/magazine/wikipedia-ai-chatgpt.html?smid=url-share	4
Kevin Roose, <i>How Does ChatGPT Really Work?</i> , N.Y. Times (Apr. 4, 2023), https://www.nytimes.com/2023/03/28/technology/ai-chatbots-chatgpt-bing-bard-llm.html	4, 19
The New York Times Company, 2023 Annual Report at 14, https://nytimes.com/2024/03/2023-Annual-Report_WR_Final.pdf	4
<i>New York Times' Granular Gen AI Tool Drives Campaign Performance</i> , Adweek (July 24, 2024), https://www.adweek.com/media/exclusive-how-the-new-york-times-granular-gen-ai-tool-drives-campaign-performance/	3
Press Release: New York Times Advertising launches BrandMatch Out of Beta (July 24, 2024), https://www.nytimes.com/press/new-york-times-advertising-launches-brandmatch-out-of-beta/	3
Press Release: Principles for Using Generative A.I. in The Times's Newsroom, N.Y. Times (May 9, 2024), https://www.nytimes.com/press/principles-for-using-generative-ai-in-the-times-newsroom/	3
Press Release: Zach Seward Is the Newsroom's Editorial Director of A.I. Initiatives, N.Y. Times (Dec. 12, 2023), https://www.nytimes.com/press/zach-seward-is-the-newsrooms-editorial-director-of-a-i-initiatives/	2
Sara Fischer, <i>Exclusive: NYT Plans to Debut New Generative AI Ad Tool Later This Year</i> , Axios (Feb. 20, 2024), https://www.axios.com/2024/02/20/new-york-times-ai-ad-tool	3
Steven Johnson, <i>A.I. Is Mastering Language. Should We Trust What It Says?</i> , The New York Times Magazine (Apr. 15, 2022), available at https://www.nytimes.com/2022/04/15/magazine/ai-language.html?smid=url-share	4

I. INTRODUCTION

This is an objection under Fed. R. Civ. Proc. 72(a) seeking to “modify or set aside” a discovery order issued by the Magistrate Judge in this case. Defendant OpenAI, a leading developer of “generative AI” technology, has sought from Plaintiff The New York Times Company documents and other information regarding the Times’s own use and development of generative AI. The Times has refused to provide it, contending that it can demand billions in damages against OpenAI without revealing anything at all about the Times’s own use, development, or analysis of the exact same technology.

That position, which the order challenged here embraced, is categorically wrong. It is wrong as a matter of legal doctrine: the fair use inquiry, which is among the primary defenses to alleged liability in this litigation, permits discovery into a plaintiff’s own use of the exact same technology that it contends in litigation somehow imperils the very enterprise upon which it is embarked. *Cf.* ECF 170 (“FAC”) ¶ 2 (asserting that OpenAI’s AI products “threaten[] The Times’s ability to provide” high-quality journalism). But it is also wrong as a matter of common sense: it is unquestionably relevant to this case whether the Times is speaking out of both sides of its mouth—at once denouncing OpenAI’s technology at every turn, even as the Times itself has deployed and even developed the same kinds of tools, using the same kinds of development methods, that it seeks to persuade the Court are unlawful.

OpenAI respectfully seeks the Court’s intervention to ensure the Times does not shield from discovery crucial information about its own use and development of generative AI tools, as it has claimed entitlement to do to date.

II. BACKGROUND

A. The Times's Contradictory Allegations and Actions

The New York Times filed this lawsuit on December 27, 2023, against Microsoft Corporation and various OpenAI entities. ECF 1. The Times's core allegation is that OpenAI used a “copy of the Internet” to train the large language models that underlie its ChatGPT service—and that, in so doing, OpenAI created unauthorized copies of Times articles published on the Internet. FAC ¶¶ 2, 88. The Times's central legal theory is that the creation of these copies constituted infringement of copyright, rather than fair use.

Despite pursuing that legal theory in Court, the Times's own public statements and actions indicate that it, too, has either created or deployed in its business large language models trained on a “copy of the Internet”—including numerous works it neither owns nor licensed from other publishers. The Times, for example, is a longtime investor in, and “early customer” of, at least one “copywriting software company” whose services rely on large language models created by OpenAI and other AI companies.¹ Shortly after filing this lawsuit, the Times announced an “A.I. Initiatives Team” tasked with “determin[ing] where to incorporate generative A.I. tools into our publishing.”² The Times then began to explicitly disclose its use of “several large-language models” for reporting purposes.³ And the Times has loudly touted its creation of a product called

¹ Carl Franzen, *Marketing Software Anyword Adds Integrations to ChatGPT, Notion AI, and Canva AI*, Venture Beat (July 18, 2023), <https://venturebeat.com/ai/marketing-software-anyword-adds-integrations-to-chatgpt-notion-ai-and-canva-ai/>.

² Press Release: Zach Seward Is the Newsroom's Editorial Director of A.I. Initiatives, N.Y. Times (Dec. 12, 2023), <https://www.nytimes.com/press/zach-seward-is-the-newsrooms-editorial-director-of-a-i-initiatives/>.

³ Alexandra Berzon, et al., *Inside the Movement Behind Trump's Election Lies*, N.Y. Times (Oct. 28, 2024), <https://www.nytimes.com/interactive/2024/10/28/us/politics/inside-the-movement-behind-trumps-election-lies.html>.

BrandMatch, “its GenAI-powered ad targeting solution”⁴ that, according to the Times, has “contributed to the success of [the Times’s] ad business.”⁵ The Times disclosed that, in developing that product, it had “experiment[ed] with multiple large language models (LLMs), including [both] enterprise models and open-source models,” which it intended to use to “power the new tool.”⁶ More recent reporting has said that the released version of the BrandMatch product, which the Times “has been using” since April 2024, “uses large language models”—although the Times has generally declined to “share specifics.”⁷ That has been a consistent theme: when the Times published a set of “Principles” to govern its use of AI, that document conspicuously failed to mention whether the large language models the Times uses were trained on unlicensed data—and similarly omitted any commitment to license the training data it uses for its own models (or to use models from developers who did so).⁸

If the Times has indeed deployed “large-language models” in its business operations—as it has publicly admitted, *see supra* 3—there is every reason to believe that those models were trained, whether by the Times or others, on large amounts of unlicensed data collected from the Internet, including news articles owned by other publishers. As the Times has explained, “[t]o get a large enough data set to make a [large language model] work, you need to scrape the wider

⁴ Press Release: New York Times Advertising launches BrandMatch Out of Beta (July 24, 2024), <https://www.nytc.com/press/new-york-times-advertising-launches-brandmatch-out-of-beta/>.

⁵ Trishla Ostwal, *Exclusive: How the New York Times’ Granular Gen AI Tool Drives Campaign Performance*, Adweek (July 24, 2024), <https://www.adweek.com/media/exclusive-how-the-new-york-times-granular-gen-ai-tool-drives-campaign-performance/>.

⁶ Sara Fischer, *Exclusive: NYT Plans to Debut New Generative AI Ad Tool Later This Year*, Axios (Feb. 20, 2024), <https://www.axios.com/2024/02/20/new-york-times-ai-ad-tool>.

⁷ *See supra* note 5.

⁸ Press Release: Principles for Using Generative A.I. in The Times’s Newsroom, N.Y. Times (May 9, 2024), <https://www.nytc.com/press/principles-for-using-generative-aie2%80%a4i%e2%80%a4-in-the-timess-newsroom/>.

web.”⁹ The Times’s own reporting acknowledges that collecting a “colossally large repository of text”—typically compiled by collecting “billions of pages . . . from the internet”—is one of the key steps to creating a modern large language model.¹⁰ Indeed, numerous third-party creators of large language models are, like OpenAI, currently embroiled in similar litigation regarding whether the creation of such models is covered by fair use.¹¹

That is presumably why the Times recently cautioned investors about its “use” of “certain generative AI tools in [its] business”—which it saw fit to classify as a “risk factor[]” in its recent Annual Report.¹² In so doing, the Times specifically noted that the technology “presents emerging legal . . . issues”—referring to the ongoing fair-use litigation against OpenAI and other companies, *see supra* note 11—and warned investors that, as a result of these issues, its “use” of these (undisclosed) “generative AI tools” may “adversely affect our business, brand, financial condition, or results of operations.”¹³ In other words, the Times warned investors that the Times *itself* could

⁹ Steven Johnson, *A.I. Is Mastering Language. Should We Trust What It Says?*, The New York Times Magazine (Apr. 15, 2022), available at <https://www.nytimes.com/2022/04/15/magazine/ai-language.html?smid=url-share>; *see also* Jon Gertner, *Wikipedia’s Moment of Truth*, The New York Times Magazine (July 18, 2023), available at <https://www.nytimes.com/2023/07/18/magazine/wikipedia-ai-chatgpt.html?smid=url-share> (“For years it has been clear that fledgling A.I. systems were being trained . . . [via a] process whereby engineers ‘scrape’ the web to create enormous data sets for that purpose”).

¹⁰ Kevin Roose, *How Does ChatGPT Really Work?*, N.Y. Times (Apr. 4, 2023), <https://www.nytimes.com/2023/03/28/technology/ai-chatbots-chatgpt-bing-bard-llm.html?smid=url-share>.

¹¹ *See, e.g., Kadrey v. Meta Platforms*, No. 23-cv-03417 (N.D. Cal.); *J.L. v. Alphabet Inc.*, No. 23-cv-03440 (N.D. Cal.); *Concord Music Group, Inc. v. Anthropic PBC*, No. 24-cv-03811 (N.D. Cal.); *Nazemian v. Nvidia Corp.*, No. 23-cv-01454 (N.D. Cal.); *O’Nan v. Databricks, Inc.*, No. 24-cv-01451 (N.D. Cal.); *Dubus v. Nvidia Corp.*, No. 24-cv-02655 (N.D. Cal.); *Makkai v. Databricks, Inc.*, No. 3:24-cv-02653 (N.D. Cal.); *Zhang v. Google LLC*, No. 24-cv-02531 (N.D. Cal.); *Millette v. Google LLC*, No. 24-cv-04708 (N.D. Cal.); *Millette v. Nvidia Corp.*, No. 24-cv-05157 (N.D. Cal.).

¹² The New York Times Company, 2023 Annual Report at 14, https://nytcassets.nytimes.com/2024/03/2023-Annual-Report_WR_-Final.pdf.

¹³ *See supra* note 12.

face legal risk from the very issue the Times is teeing up *against OpenAI* in this case: whether the use of textual works to train large language models is fair use.

B. OpenAI Hewed Closely to Precedent When Seeking Evidence

Because this is a fair use case, its merits depend on a “multifaceted assessment” of whether the challenged use serves the goals of copyright. *Authors Guild v. Google, Inc. (Google Books)*, 804 F.3d 202, 213 (2d Cir. 2015); *id.* (fair use asks whether the use “serve[s] the overall objective of the copyright law to expand public learning”). To that end, a fair use defense often involves consideration of whether the *copyright holder’s* own words or deeds support a finding of fair use. The Supreme Court’s decision in *Google LLC v. Oracle America Inc.* is instructive. 593 U.S. 1 (2021). In that case, Oracle (the copyright holder) sued Google (the alleged infringer) for “reimplement[ing]” elements of the “interface” of Oracle’s Java programming language, and Google claimed fair use as a defense. *Id.* at 31–32. In finding fair use as a matter of law, the Court specifically highlighted evidence that Oracle had—through words and deeds—endorsed the same practice of “reimplementing [software] interfaces” that it challenged as unlawful in the litigation. *Id.* (explaining that “[t]he jury heard” that Oracle “itself had used pre-existing interfaces in creating Java” and noting that Oracle’s general practices included the “reimplementation of other APIs”). That evidence was important to “convinc[ing]” the Court that “Google’s copying was transformative”—a key element in the fair use analysis. *Id.* The Court also examined evidence of Oracle’s “efforts to move into the mobile phone market,” which “had proved unsuccessful.” *Id.* at 36. That evidence, in turn, supported the Court’s conclusion that the “fourth factor—market effects—also weigh[ed] in favor of fair use.” *Id.* at 40.

To make the same showing—and because of the numerous public statements suggesting the Times too has embraced the very technological practice it challenges in litigation, *see supra* 2–4—OpenAI served discovery requests to investigate the Times’s creation, use, and adoption of

large language models in its business. ECF 236-1 (OpenAI First Set of RFPs). OpenAI tailored its requests to target the types of evidence courts consider when addressing fair use—specifically, evidence of the plaintiff’s own activities that go to the transformativeness of the use and the absence of market harm. *See supra* 5 (discussing *Oracle*); *see also Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 930 (2d Cir. 1994) (considering evidence of industry practices in analyzing the fourth fair use factor).

OpenAI thus sought from the Times the very types of evidence the Supreme Court found relevant in *Oracle*: (1) documents concerning the Times’s own “adoption of Generative AI for use by [its] Employees,” *see, e.g.*, ECF 236-1 at 15 (RFP 45); *see also, e.g., id.* at 14–15 (RFPs 44, 46, 49); (2) documents concerning its “past and present efforts or future plans to utilize Generative AI in [its] reporting and presentation of content,” *see, e.g., id.* at 15 (RFP 49); and (3) documents “relating to the training” of various generative AI tools the Times had disclosed in public statements, *see, e.g., id.* at 15–16 (RFPs 50, 51). OpenAI also sought the Times’s communications with certain third-parties about “copyright and artificial intelligence,” a category that would include statements by the Times regarding the “purpose” and “character” of the use and the impact (if any) on the market for or value of the copyrighted works—also something the *Oracle* Court found relevant. *Id.* at 11 (RFP 24); *see also Oracle*, 593 U.S. at 31 (relying on evidence that Oracle’s “executives thought that widespread use of the Java programming language” would inure to Oracle’s benefit).

The Times refuses to comply with those requests, insisting that it is entitled to keep secret its own development and use of non-OpenAI large language models. More specifically, the Times refuses to agree to produce documents addressing its use of any non-OpenAI large language models, which would include models such as “Gemini” (a model created by Google) and “Claude”

(a model created by Anthropic). ECF 236-4 at 4 (NYT’s Resp. & Obj. to OpenAI’s First Set of RFPs). The Times likewise refuses to agree to produce discovery related to its own development of large language models, including discovery about what the Times used to train those models. ECF 236-6 at 1 (arguing “use of AI generally is [] irrelevant” because “[t]his case is about Defendants’ GenAI tools”). Simply put, the Times believes that OpenAI—and the public—are not entitled to know whether the Times itself has developed or is developing large language models using the exact same methods that the Times impugns as illegal in this litigation.

What the Times has agreed to produce related to its use of non-OpenAI generative AI is a narrow subset of documents related to its “A.I. Initiatives” program. ECF 236-6 at 1. And while the Times has produced a handful of such documents, none provides information about the Times’s training of large language models—let alone details about the non-OpenAI models the Times uses in its business operations.

C. The Order Denied OpenAI’s Request by Ignoring Governing Law

On September 3, 2024, OpenAI sought a pre-motion conference before Magistrate Judge Ona T. Wang to address the Times’s refusal to comply with these requests. ECF 236. On November 22, 2024, the court denied OpenAI’s motion without full briefing and without permitting OpenAI to file a motion to compel. *See generally* ECF 344 (“Order”).

The court concluded that—as a matter of law—the discovery OpenAI seeks is not relevant to its fair use defense. According to the court, the four statutory fair use factors only contemplate scrutiny of the *defendant’s* actions and thus “do not require a court to examine” the plaintiff’s related statements or actions. Order at 2–3 (emphasis in original). Based on that analysis, the court concluded that the evidence OpenAI sought regarding the plaintiff’s statements and actions are categorically irrelevant to the fair use determination. *Id.* In reaching that conclusion, the court did not address the fact that—as the Supreme Court has cautioned—the “list of factors is not

exhaustive,” *Oracle*, 593 U.S. at 19, and the statute is intended to “provide only *general guidance* about the sorts of copying that courts and Congress most commonly had found to be fair uses.” *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 577–78 (1994) (emphasis added); *see also Oracle*, 593 U.S. at 31–32 (looking to copyright holder’s activities and its executives’ statements in evaluating fair use).

The court then discussed the fair use rulings in *Oracle, Texaco*, 60 F.3d at 929–30, and *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023). In the court’s view, none supported OpenAI’s position that evidence OpenAI sought regarding the Times’s own activities related to generative AI is relevant to OpenAI’s fair use defense. *See* Order at 3–5. Instead, according to the court, fair use looks at what the defendant has or has not done—rendering almost all plaintiff-related evidence irrelevant and undiscoverable. *See id.*; *contra Texaco*, 60 F.3d at 927–29 (relying on evidence relating to activities of “many major corporations” other than the defendant). The court did not address the fact that *Oracle* considered the very type of evidence OpenAI sought here. *Oracle*, 593 U.S. at 31–32 (citing evidence that plaintiff had “reimplement[ed] other APIs”); *id.* at 36 (analyzing “evidence” of plaintiff’s “many efforts to move into the mobile phone market” which “proved unsuccessful”).

The court also waved away the very premise of OpenAI’s requests by assuming—without any factual basis—that the Times simply could not have acted in a manner that contradicted its litigation position. It rejected the notion that the Times’s own use of third-party generative AI products could be relevant because, in the court’s view, any services the Times may have used were “presumably . . . developed without copying.” Order at 4. The court did not cite any support for that presumption, and all indications suggest it is false: indeed, not even the Times has claimed that the large language models it uses were “developed without copying.” *Id.* That is not

surprising since, according to the Times, “[t]o get a large enough data set to make a[] [large language model] work, you need to scrape the wider web.” *See supra* note 9.

III. STANDARD OF REVIEW

Under Federal Rule of Civil Procedure 72(a), a district court may “modify or set aside” any part of a magistrate judge’s non-dispositive order “that is clearly erroneous or is contrary to law.” Fed. R. Civ. P. 72(a). A magistrate judge’s order is “contrary to law” if it “fails to apply or misapplies relevant statutes, case law, or rules of procedure.” *Coventry Capital U.S. LLC v. EEA Life Settlements Inc.*, 333 F.R.D. 60, 64 (S.D.N.Y. 2019) (quoting *R.F.M.A.S., Inc. v. So*, 748 F. Supp. 2d 244, 248 (S.D.N.Y. 2010)). A magistrate judge’s order is “clearly erroneous” “if the reviewing court is left with the definite and firm conviction that a mistake has been committed.” *Id.* (quoting *R.F.M.A.S., Inc.*, 748 F. Supp. 2d at 248).

IV. ARGUMENT

Federal Rule of Civil Procedure 26(b)(1) provides that parties “may obtain discovery regarding any nonprivileged matter that is relevant to any party’s claim or defense[.]” “Although not unlimited, relevance, for purposes of discovery, is an extremely broad concept.” *Chen-Oster v. Goldman, Sachs & Co.*, 293 F.R.D. 557, 561 (S.D.N.Y. 2013) (quoting *Condit v. Dunne*, 225 F.R.D. 100, 105 (S.D.N.Y. 2004)). It is thus “well-settled within this Circuit that [Rule 26] will be satisfied if there is ‘any possibility that the information sought may be relevant to the subject matter of the action.’” *Forte v. Liquidnet Holdings, Inc.*, No. 14-cv-2185, 2014 WL 12775220, at *1 (S.D.N.Y. Dec. 2, 2014) (emphasis added) (quoting *Daval Steel Prod., a Div. of Francosteel Corp. v. M/V Fakredine*, 951 F.2d 1357, 1367 (2d Cir. 1991)). Here, the court concluded that the discovery OpenAI is seeking does not fall within the “extremely broad concept” of relevance. *Chen-Oster*, 293 F.R.D. at 561. That conclusion is both clearly erroneous and contrary to law.

A. The Order is Contrary to Binding Fair Use Precedent

The doctrine of fair use—an “equitable rule of reason” that “originat[ed] in the courts”—“permits courts to avoid rigid application of the copyright statute when, on occasion, it would stifle the very creativity which that law is designed to foster.” *Oracle*, 539 U.S. at 18 (quoting *Stewart v. Abend*, 495 U.S. 207, 236 (1990)). Fair use “guarantee[s] [] breathing space within the confines of copyright” by excusing acts of copying that further copyright’s goal—“to promote science and the arts”—without unduly undercutting copyright’s core incentive to create. *Campbell*, 510 U.S. at 578. Congress codified fair use in 17 U.S.C. § 107, where it provided a list of factors for consideration when evaluating a fair use defense:

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- (2) the nature of the copyrighted work;
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- (4) the effect of the use upon the potential market for or value of the copyrighted work.

17 U.S.C. § 107. The flexible nature of the fair use defense, along with the baseline rule that relevance under the Federal Rules is “broad and liberally construed,” *Bamonte v. Charatan*, No. 22-cv-0795, 2023 WL 4201416, at *3 (S.D.N.Y. June 27, 2023), means that the discovery sought here is highly relevant and should be produced.

The Order misapprehends the fair use inquiry in multiple respects. First, the Order suggests that the statutory language of Section 107 limits the categories of evidence that are relevant to fair use—and, in particular, precludes consideration of the key statements, conduct, and practices by the plaintiff. Order at 2–3. That ruling is irreconcilable with Supreme Court precedent. *Infra* Section IV.A.1. Second, the Order ignores that the evidence OpenAI seeks is relevant to the

existence (or non-existence) of a workable market through which OpenAI (or the Times) could have licensed training data for this purpose—an inquiry that goes directly to the fourth fair use factor. *Infra* Section IV.A.2. And third, the Order ignores the relevance of this evidence to the question of whether the use at issue here is transformative (it is) and whether it yields precisely the kinds of public benefits that copyright exists to promote (it does)—issues that are relevant to the first and fourth fair use factors. *Infra* Section IV.A.3.

1. The Order Erroneously Concludes that Plaintiff’s Conduct is Almost Entirely Irrelevant to Fair Use

The Order holds that the statutory factors in Section 107, although “non-exclusive,” do not “require[] scrutiny” of the plaintiff’s own conduct as part of the fair use inquiry—and that the discovery OpenAI sought into that conduct is thus categorically irrelevant and undiscoverable. Order at 2–3. That category of supposedly irrelevant evidence includes, according to the Order, “whether the copyright holder has used tools” similar to those it challenges in litigation or “admi[ssions]” that might bear on a defendant’s arguments. *Id.*

That conclusion is contrary to law. As explained above and below, the Supreme Court explicitly relied on similar evidence in *Oracle*. This evidence was relevant to multiple statutory factors, including factor 1 and factor 4. Moreover, as the Court explained, Section 107’s “list of factors is not exhaustive,” *Oracle*, 593 U.S. at 19, and the prefatory list of fair use examples is “illustrative and *not limitative*,” providing “only *general guidance* about the sorts of copying that courts and Congress most commonly had found to be fair uses.” *Campbell*, 510 U.S. at 577–78 (emphasis added); *id.* (statute not intended to “change” or “narrow” the “judicial doctrine of fair use”). As such, fair use analysis often involves consideration of evidence that bears on issues beyond the statutory language in Section 107: the doctrine, properly construed, requires a “multifaceted assessment.” *Google Books*, 804 F.3d at 213; *see also Oracle*, 593 U.S. at 35 (noting

fourth factor is “more complex than at first it may seem”).

That is one of the many reasons why the range of relevant discovery in fair use cases necessarily includes evidence relating to the *plaintiff*’s activities—including efforts by a plaintiff to enter a particular market, *see, e.g., Oracle*, 593 U.S. at 36, and documents from the plaintiff’s files that suggest acknowledgement that the challenged conduct is “noninfringing” and “d[id] not require a license,” *see, e.g., Pearson Educ., Inc. v. Chegg, Inc.*, No. 21-cv-16866, 2023 WL 3775317, at *1 (D.N.J. June 2, 2023) (granting motion to compel). Where, for example, a plaintiff’s business activities appear to contradict its legal position with respect to fair use, as was the case in *Chegg*, “discovery into whether [the plaintiff] believed that [the defendant’s actions] were non-infringing . . . is directly relevant to [the defendant’s] fair use defense.” 2023 WL 3775317, at *7. And in other areas of law, courts do not hesitate to find relevant evidence that a plaintiff has made statements or undertaken activity that directly conflicts with its position in litigation. *See, e.g., Freedom Card, Inc. v. JPMorgan Chase & Co.*, 432 F.3d 463, 476 (3d Cir. 2005) (holding that plaintiff’s “prior representations” about the enforceability of a trademark “undermine [its] attempt to establish” trademark liability).¹⁴

Indeed, and as noted above, this category of evidence took center stage in the *Oracle* litigation. As Justice Breyer explained:

Some of the *amici* refer to what Google did as “reimplementation” The record here demonstrates the numerous ways in which reimplementing an interface can further the development of computer programs. The jury . . . heard that the reuse of APIs is common in the industry. It heard that [Oracle] *itself had used pre-existing interfaces* in creating Java. And it heard that Sun executives thought that widespread use of the Java programming language . . . would benefit the company.

¹⁴ *See also Int’l Bus. Machs. Corp. v. Naganayagam*, No. 15-cv-7991, 2016 WL 11483935, at *4 (S.D.N.Y. Dec. 9, 2016) (in contract dispute that turned on whether IBM and a non-party were “competitors,” holding that “defendant is entitled to” IBM’s “internal documents” that might contradict its litigation position).

Oracle, 593 U.S. at 31–32 (emphasis added) (citations omitted). *None* of this evidence would have been discoverable if—as the Order below holds—courts cannot consider as part of a fair use analysis “comments a [plaintiff] may have made about a defendant’s general industry,” Order at 2–3, or a plaintiff’s own use of the business practices it challenges in litigation, *id.* See *Oracle*, 593 U.S. at 31–32 (citing to amicus brief regarding “Oracle’s reimplementation of other APIs” to support the point that this conduct “can further the development” of the relevant industry).

By cabining the scope of fair use discovery in a manner contrary to binding precedent—*i.e.*, to exclude evidence relating to the Times’s own statements and conduct—the ruling threatens to limit the availability of critical and relevant evidence for summary judgment, trial, and appeal. This Court should sustain the objection to correct this core legal error.

2. The Order Misconstrues the Market Harm Inquiry

The Order also addresses the types of discovery the court views as relevant to the fourth fair use factor—“the effect of the use upon the potential market for or value of the copyrighted work,” 17 U.S.C. § 107(4). This factor looks at, as relevant here, “whether there was harm to the actual or potential markets for the copyrighted work.” *Oracle*, 593 U.S. at 26 (quoting *Oracle Am., Inc. v. Google LLC*, 886 F.3d 1179, 1196 (Fed. Cir. 2018)). That inquiry involves consideration of, *inter alia*, whether there exists a “traditional, reasonable, or likely to be developed market[]” for licensing the use in question. *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 614 (2d Cir. 2006) (quoting *Texaco*, 60 F.3d at 930).

Evidence that the Times has itself created or deployed a large language model trained on unlicensed data bears directly on whether a “ready market or means to pay for the use” exists. *Bill Graham*, 448 F.3d at 416. Given the Times’s litigation position—including its insistence that model developers “should first approach [rightsholders] about a licensing agreement,” FAC ¶ 49—it is reasonable to expect that the Times evaluated whether licensing would be feasible before

deciding that it was not. In so doing, the Times would have evaluated, for example, the existence of licensing mechanisms, the cost and economics of those mechanisms, the volume of training data that it would need to license, and the transaction costs of securing such licenses.

All such evidence is directly relevant to the question whether there is a “ready market or means” through which OpenAI (or the Times) could have “pa[id] for the use” of sufficient data to train large language models. *Texaco*, 60 F.3d at 929–31. The same is true for the use of large language models created by others using unlicensed data. The Times’s decision to use such models *knowing* that they were trained on unlicensed data would suggest that the Times had evaluated the possibility of finding a large language model provider who obtained license agreements before training—and rejected that possibility as unworkable. *See supra* 3–4 (noting that all useful large language models use unlicensed data).

Put differently, if a functional market for licensing training data existed, the Times would have participated in it—either by obtaining licenses for the data it used to train its own models or engaging exclusively with third-party model developers who did the same. Evidence that the Times did not—and that the “several large-language models” it has deployed were created by the Times or others with unlicensed data, *see supra* note 3—goes directly to whether such a market is even possible let alone actually exists.

According to the Order, however, what matters for purposes of the fourth factor are “documents concerning licensing discussions” and “discovery from [OpenAI] on how its use might ‘kill demand for the original.’” Order at 4 (quoting *Oracle*, 593 U.S. at 35). Evidence of the plaintiff’s “other use or licensing of their own works to other nonparties,” the court concluded, is “not at issue in the fair use determination.” Order at 5. And the court did not even discuss the possible relevance of the Times’s adoption of large language models to the question whether there

exists a “ready market” for training data. *See generally* Order.

This holding is contrary to binding Supreme Court and Second Circuit precedent. The Supreme Court in *Oracle* and the Second Circuit in *Texaco* both explicitly relied on market evidence from parties *other than* the defendant—including evidence relating to the plaintiff’s activity in the market—to evaluate the fourth factor. *See Oracle*, 593 U.S. at 36 (analyzing “evidence” of the plaintiff’s “many efforts to move into the mobile phone market” which “provided unsuccessful”); *Texaco*, 60 F.3d at 930 (noting that “[t]he District Court found that many major corporations now subscribe” to “systems for photocopying licenses,” which suggests that “there currently exists a viable market for licensing these rights”).

And the Times itself has effectively (and rightly) conceded the relevance of market-related evidence from parties other than the defendant. It has served subpoenas on numerous third-party developers of large language models, seeking discovery into those parties’ market-related activity, such as their “[l]icensing agreements with content owners” related to their generative AI products. *See, e.g.*, Ex. D (subpoena to Anthropic). The Times should not be permitted to withhold evidence of its own assessment of the existence or absence of a market for licensing content for generative AI products while at the same time demanding market-related evidence from third parties.

The Order, in other words, ignores this major element of fourth-factor analysis, finding that the relevant evidence a plaintiff may have would be reflected in “documents concerning licensing discussions.” Order at 4. That narrow view of the plaintiff-related evidence relevant to the fourth fair use factor is contrary to law and threatens to hamstring OpenAI’s ability to mount its fair use defense. *See* Fed. R. Civ. P. 72(a) (“The district judge in the case must . . . modify or set aside any part of the order that is . . . contrary to law.”).

3. The Order Overlooks the Relevance to Transformativeness and Public Benefit

Fair use—and, in particular, the first and fourth factors—also considers whether the use advances the goal of copyright by promoting science and the arts. The first factor functions as the “heart of the fair use inquiry,” *On Davis v. The Gap, Inc.*, 246 F.3d 152, 174 (2d Cir. 2001) (Leval, J.), and often involves analysis of the technology at issue within the broader legal framework—including by looking at whether the technology “serve[s] the overall objective[] of the copyright law to expand public learning.” *Google Books*, 804 F.3d at 213; *see also id.* at 218 (first factor satisfied in part because service has a “highly transformative purpose of identifying books of interest to the searcher”).¹⁵ Similarly, courts “take into account the public benefits the copying will likely produce” when analyzing the fourth factor, including by asking whether those benefits are “related to copyright’s concern for the creative production of new expression[.]” *Oracle*, 593 U.S. at 35–36. The court’s decision is contrary to those decisions as well.

Under these precedents, the fair-use inquiry considers the extent to which the copying serves “copyright’s very purpose, to promote the Progress of Science and the useful Arts.” *Campbell*, 510 U.S. at 575 (quoting U.S. CONST. Art. 1, § 8, cl. 8).¹⁶ OpenAI intends to show that, by developing its large language model, it has unlocked a virtually infinite array of new and unprecedented capabilities that directly promote the “Progress of Science and the useful Arts.” U.S. CONST. Art. 1, § 8, cl. 8 (Patent and Copyright Clause). Evidence that the Times itself has used the same technology to improve and further its journalistic enterprise would directly support

¹⁵ *See also Sega Enterprises Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1523 (9th Cir. 1992) (first factor satisfied where use “led to an increase in the number of independently designed video game programs” which is “precisely [the kind of] growth in creative expression, based on the dissemination of other creative works and the unprotected ideas contained in those works, that the Copyright Act was intended to promote”).

¹⁶ ECF 52 at 4–8 (describing OpenAI’s development of this technology).

OpenAI’s position. *See Compudyne Corp. v. Shane*, 244 F.R.D. 282, 282–83 (S.D.N.Y. 2007) (evidence that would undercut plaintiff’s theory of the case and could potentially impeach witnesses was relevant); *Google Books*, 804 F.3d at 217 (evidence that technology enabled “readers to learn the frequency of usage of selected words in the aggregate corpus of published books in different historical periods” “strongly favor[ed] satisfaction of the first factor”); *Oracle*, 593 U.S. at 39 (noting, in the fourth factor analysis, that rejecting the fair use defense in this context would “interfere with, not further, copyright’s basic creativity objectives” by discouraging the creation of “creative improvements, new applications, and new uses”).

The Times implicitly conceded this point by acknowledging that its use of models *created by OpenAI and Microsoft* is relevant and discoverable. *See* Order at 1 n.1. But the Order concludes that evidence of the Times’s use of *other* models is irrelevant because the FAC includes “no wholesale indictment of Gen AI tools.” *Id.* at 3; *but see* FAC at 14 (“GenAI Products Threaten High-Quality Journalism.”). But that misses the point: the Times’s use of a large language model built in the same way using the same technology has the same evidentiary value to OpenAI’s defense as the Times’s use of OpenAI’s models. The fact that the Times did not refer to these other companies in its complaint against OpenAI does not mean that OpenAI may not use this evidence to build its own defense. In any case, the Order did not cite a single case or authority suggesting that the “extremely broad concept” of relevance, *Chen-Oster*, 293 F.R.D. at 561 (quoting *Condit*, 225 F.R.D. at 105), is so limited. *See generally* Order. Nor would that limitation make any sense: if the Times has decided to employ this technology (as it has suggested, *see supra* 2–4), it is likely to have relied predominantly on models created either by itself or by companies it has *not* sued for “billions of dollars.” FAC ¶¶ 8–9.

B. The Order Shields Relevant Evidence from Discovery

The Times’s core theory in this case is that the “unlicensed use of copyrighted content to train GenAI models” like large language models is “not fair use.” FAC ¶ 8. As explained above, however, the Times’s own actions suggest that it has either implicitly or expressly relied on the very same fair use doctrine to justify its own use of large language models trained on unlicensed content—something that the Times, not surprisingly, wants to keep from the public eye.

1. The Documents OpenAI Seeks are Directly Relevant

Evidence that the Times has engaged in the conduct that it contends is improper in this lawsuit would directly undermine the Times’s claims—and support OpenAI’s defenses—in this case. If, for example, Times executives have stated internally that the company’s use of unlicensed data to create its own “large-language models” is transformative, documents evidencing those views would qualify as “[i]nconsistent statements” that contradict the Times’s argument that there is “nothing ‘transformative’” about the activity it challenges in this lawsuit—and thus may be used to impeach the testimony of those executives. *See* FAC ¶ 8; Wright & Miller, 8 FED. PRAC. & PROC. CIV. § 2015 (“Inconsistent statements . . . cannot be excluded from the scope of discovery.”).¹⁷ If the Times’s leadership has paid another AI company—like Meta, Anthropic, Nvidia, Google, or others—to use a large language model that the Times knew or suspected to have been trained on unlicensed data, documents evidencing those voluntary commercial arrangements would likewise undercut the Times’s argument that such conduct should be considered “unlawful.” FAC ¶ 154; *see also* Order at 4–7 in *Arista Records LLC, et al. v. Lime Group LLC*, No. 06-cv-5936, Dkt. 363 at 4–5 (S.D.N.Y. Nov. 19, 2010) (affirming order requiring

¹⁷ The Times has used a similar argument to justify its demand for discovery into the internal files of certain OpenAI former employees. *See, e.g.*, ECF 205 at 3.

production of documents that “could illuminate Plaintiffs’ ‘conduct and attitude’ about its copyrights, licensing, and internet companies generally” and “show how Plaintiffs conducted themselves in dealing with others in the Internet marketplace”). Such evidence is directly relevant to the Times’s claims and OpenAI’s defense in this case, and is thus discoverable.

2. The Court Relied on an Unwarranted Factual Assumption about the Times’s Own Conduct

The court ignored all of this by assuming that, if the Times *had* used large language models in its business, those models were “presumably [] developed without copying.” Order at 4. The Order, in other words, *presumes* that the activity described in the documents OpenAI is seeking is not the same kind of activity the Times challenges in this litigation as illegal.

That was clearly erroneous—not least because *even the Times* has tacitly conceded that it has created or deployed large language models trained on unlicensed data, and has even disclosed this activity as a legal risk in a regulatory filing. *See supra* 4. As discussed above, the Times has explicitly disclosed its use of “several large-language models” in business operations, *see supra* 3, and the Times has reported that large language models are “usually” created by “scrap[ing]” a “large repository of text” from the Internet.¹⁸ There is, in other words, every reason to expect that the “large-language models” that the Times has deployed in its own business were created through exactly the same methods the Times claims in this case are improper (*i.e.*, through the use of unlicensed content collected from the Internet) and internally justified its actions as permissible for the exact same reason OpenAI asserts as a defense in this litigation—that such conduct

¹⁸ Kevin Roose, *How Does ChatGPT Really Work?*, N.Y. Times (Apr. 4, 2023), <https://www.nytimes.com/2023/03/28/technology/ai-chatbots-chatgpt-bing-bard-llm.html> (“I’m going to walk you through setting up a large language model from scratch . . . [i]deally, we’ll put together a colossally large repository of text” by “us[ing] some free, publicly available data libraries, such as the Common Crawl repository of web data.”).

constitutes fair use. *Cf. Dawe v. Capital One Bank*, No. 04-cv-40192, 2005 WL 8176463, at *2 (D. Mass. Sept. 27, 2005) (“[I]nformation is discoverable if there is *any possibility* it might be relevant to the subject matter of the case.” (emphasis added)). Indeed, in the almost nine months since OpenAI propounded these discovery requests, the Times has never denied using large language models trained on unlicensed data. *See supra* 8.

The Court’s factual assumption that the Times’s conduct was effectively beyond reproach was thus not only inappropriate, *see Francis v. A & E Stores, Inc.*, No. 06-cv-1638, 2008 WL 2588851, at *2 (S.D.N.Y. June 26, 2008) (courts “do[] not resolve factual disputes” “before discovery is completed”), but also erroneous. The purpose of OpenAI’s discovery requests, as explained above, was to adduce documents regarding the Times’s use of the very technology it challenges as illegal in this litigation. The Order below effectively rejects OpenAI’s requests, in part, by assuming that those documents do not exist—and ordering that the Times need not even bother to search for them. That was clear error.

C. The Order Ignores the Relevance to Damages and Other Remedies

The evidence OpenAI seeks is also relevant to numerous other aspects of this case. For example, the Times alleges that OpenAI engaged in “willful” infringement and thus should be liable for greater damages. FAC ¶¶ 124, 165; *see* 17 U.S.C. § 504(c)(2) (setting damages amounts for infringement “committed willfully”). The Times also requested a permanent injunction against OpenAI that orders destruction of OpenAI’s language models, as well as “[a]n award of costs, expenses, and attorneys’ fees.” *See id.* Prayer for Relief ¶¶ 2–4. These remedies will require an assessment of the “objective reasonableness” of OpenAI’s conduct and its reliance on the doctrine of fair use in this context. *See, e.g., Kirtsaeng v. John Wiley & Sons, Inc.*, 579 U.S. 197, 207–08 (2016) (“objective reasonableness” of litigation position is “an important factor in assessing fee applications” under 17 U.S.C. § 505); *see also TD Bank N.A. v. Hill*, 928 F.3d 259, 283 (3d Cir.

2019) (“the degree to which the defendant reasonably believed his conduct was non-infringing” is relevant to whether a permanent injunction in a copyright case is warranted).

Evidence that the Times itself engaged in conduct that it claims in this litigation is illegal is therefore relevant for an additional reason: to establish that OpenAI’s conduct was objectively reasonable and that its reliance on the fair use doctrine was well justified. *See, e.g., Chegg*, 2023 WL 3775317, at *6–8 (ordering production of documents suggesting that the plaintiff “acknowledged that [the challenged conduct was] non-infringing” in part because such evidence “could mitigate a charge of willful infringement” and “impact the calculus of actual and statutory damages and the availability of fee-shifting”). Similarly, excluding this evidence from the case—as the Order threatens to do—would permit the Times to continue denying the reasonableness of OpenAI’s fair use position, notwithstanding the fact that it has taken the same position to justify its own conduct. The Order below ignores this skewed result and overlooks the clear and direct relevance of this evidence to the remedies the Times affirmatively chose to seek.

V. CONCLUSION

For the foregoing reasons, OpenAI respectfully requests that this Court vacate the Order and order the Times to produce the requested discovery.

Dated: December 6, 2024

Respectfully submitted,

By: /s/ Elana Nightingale Dawson

LATHAM & WATKINS LLP

Andrew M. Gass (*pro hac vice*)

andrew.gass@lw.com

Joseph R. Wetzel

joseph.wetzel@lw.com

505 Montgomery Street, Suite 2000

San Francisco, CA 94111

Telephone: 415.391.0600

Sarang V. Damle

sy.damle@lw.com

Elana Nightingale Dawson (*pro hac vice*)

elana.nightingaledawson@lw.com

Michael David

michael.david@lw.com

555 Eleventh Street, NW, Suite 1000

Washington, D.C. 20004

Telephone: 202.637.2200

Allison L. Stillman

alli.stillman@lw.com

Rachel Renee Blitzer

rachel.blitzer@lw.com

Herman Yue

herman.yue@lw.com

Luke A. Budiardjo

luke.budiardjo@lw.com

1271 Avenue of the Americas

New York, NY 10020

Telephone: 212.906.1200

By: /s/ Joseph C. Gratz

MORRISON & FOERSTER LLP

Joseph C. Gratz (*pro hac vice*)*

jgratz@mofo.com

Vera Ranieri (*pro hac vice*)

vranieri@mofo.com

Rose S. Lee (*pro hac vice*)

rlee@mofo.com

425 Market Street

San Francisco, CA 94105

Telephone: 415.258.7522

Eric K. Nikolaides
enikolaides@mofo.com
Emily C. Wood
ewood@mofo.com
250 West 55th Street
New York, NY 10019-9601
Telephone: 212.468.8000

Carolyn Homer (*pro hac vice*)
chomer@mofo.com
2100 L Street, NW Suite 900
Washington, D.C. 20038
Telephone: 202.887.1500

By: /s/ Andrew F. Dawson

KEKER, VAN NEST & PETERS LLP

Robert A. Van Nest (*pro hac vice*)
rvannest@keker.com
Paven Malhotra
pmalhotra@keker.com
Michelle S. Ybarra (*pro hac vice*)
mybarra@keker.com
Nicholas S. Goldberg (*pro hac vice*)
ngoldberg@keker.com
Thomas E. Gorman (*pro hac vice*)
tgorman@keker.com
Katie Lynn Joyce (*pro hac vice*)
kjoyce@keker.com
Christopher S. Sun (*pro hac vice*)
csun@keker.com
R. James Slaughter (*pro hac vice*)
rslaughter@keker.com
Edward A. Bayley (*pro hac vice*)
ebayley@keker.com
Andrew F. Dawson (*pro hac vice*)*
adawson@keker.com
Andrew S. Bruns (*pro hac vice*)
abruns@keker.com
633 Battery Street
San Francisco, CA 94111-1809
Telephone: 415.391.5400

Attorneys for OpenAI

*All parties whose electronic signatures are included herein have consented to the filing of this document in accordance with Rule 8.5(b) of the Court's ECF Rules and Instructions.